

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 19, 2003, 15:21:35 ; Search time 36 Seconds

(without alignments)  
552.801 Million cell updates/sec

Title: US-09-494-297-2

Perfect score: 3945

Sequence: 1 MKKTRFPNKLTNLTQRYLS.....IAGISLGIMGHTIRIRKHD 757

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 125328 seqs, 26289065 residues

Total number of hits satisfying chosen parameters: 125328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PC1\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	149	3.8	1315	5	US-09-200-650E-5
2	142	3.6	930	6	US-10-615-383-10
3	134.5	3.4	1769	6	US-10-617-320-3676
4	134	3.4	560	6	US-10-333-002-19
5	131.5	3.3	1531	1	PCT-US02-37235-44
6	131	3.3	1456	1	PCT-US03-22113-30
7	130	3.3	5198	6	US-10-346-863-34
8	129	3.3	461	7	US-60-490-890-1651
9	128	3.2	993	6	US-10-286-897-2582
10	128	3.2	993	6	US-10-258-898A-2576
11	128	3.2	1090	6	US-10-408-765A-2576
12	128	3.2	1375	1	PCT-US03-06962-35
13	126	3.2	1342	1	PCT-US03-22113-3
14	125.5	3.2	930	5	US-09-200-650E-3
15	125.5	3.2	1972	6	US-10-617-320-3251
16	124	3.1	638	5	US-09-674-546A-2161
17	124	3.1	867	5	US-09-674-546A-2167
18	124	3.1	2142	6	US-10-603-108-3459
19	123.5	3.1	1358	6	US-10-326-956-539
20	123	3.1	980	6	US-10-326-956-539
21	122.5	3.1	437	6	US-10-163-106B-32
22	122	3.1	702	6	US-10-326-956-2663
23	121.5	3.1	640	6	US-10-603-113-1883
24	121.5	3.1	3509	6	US-10-273-573-8544
25	120.5	3.1	810	6	US-10-603-113-20281
26	120.5	3.1	903	6	US-10-333-002-10

27	120	3.0	731	6	US-10-603-113-18075	Sequence 18075, A
28	120	3.0	803	6	US-10-603-114-4886	Sequence 4886, Ap
29	120	3.0	951	7	US-60-490-890-1772	Sequence 1772, Ap
30	120	3.0	1312	6	US-10-286-897-2112	Sequence 2112, Ap
31	120	3.0	1312	6	US-10-258-898A-2112	Sequence 2112, Ap
32	120	3.0	1554	1	PCT-US03-06962-38	Sequence 38, App
33	120	3.0	1899	1	PCT-US02-39397-135	Sequence 135, App
34	120	3.0	1899	1	PCT-US02-39397-132	Sequence 132, App
35	120	3.0	2662	6	US-10-273-573-8548	Sequence 8548, Ap
36	119.5	3.0	581	6	US-10-603-113-15405	Sequence 15405, A
37	119	3.0	527	6	US-10-613-520-2128	Sequence 15917, A
38	119	3.0	1019	6	US-10-273-573-8550	Sequence 2128, Ap
39	119	3.0	2986	6	US-10-273-573-8550	Sequence 8550, Ap
40	119	3.0	4766	6	US-10-273-573-8539	Sequence 8539, Ap
41	119	3.0	4768	6	US-10-273-573-8549	Sequence 8549, Ap
42	119	3.0	5367	6	US-10-273-573-8547	Sequence 8547, Ap
43	118.5	3.0	796	6	US-10-326-956-1821	Sequence 1821, Ap
44	118.5	3.0	1347	6	US-10-408-765A-1228	Sequence 1228, Ap
45	118.5	3.0	1742	6	US-10-615-383-4	Sequence 4, Appl1

## ALIGNMENTS

## RESULT 1

US-09-200-650E-5

Sequence 5, Application US/09200650E

GENERAL INFORMATION:

Query Match 3.8% Score 149; DB 5; Length 1315;  
Best Local Similarity 20.2% Pred. No. 0.02; Mismatches 336; Indels 372; Gaps 45;  
Matches 203; Conservative 93

QY	8	NKLTNLTQRYLSKNS--KRTVTLVGYFLMFLVTSVMGAKYFGVESSTP-----	59
DB	208	NSNNNNNDIILPKTAPKRLNTR-----MRIAQAQPSSTKANNNDLITSFTLLTVDA	262
QY	60	---NAINP--DSSSEYRWYGESYRGHPYKQFVAHDLRVNLEGSNSYQYCFNLKA	114
DB	263	DKNNKIYPAQDYLSLKSQITVDDKVKSDYF-----TIKYSPTVOYQGLNPEDI	311
QY	115	FPILGS-----DSSVKWKYKHHG-----ISTKFPDYAMSPRTGDELNQLRAVMYGH	163
DB	312	KNIGIKPPNNGETATATKHDANNLLITYTFDY-----DRFNSVQMGINYSIYMDA	364
QY	164	---PONANGIMEGLPLAIRTQCAVWYSSDNAPISNPDSFKRESNL-----	211
DB	365	DTIPSKNDVENVITIGMTTKTANIQYPOVVAEKSIGSAFTIEYSHVGNKKNPGY	424
QY	212	---VSTQSLSLMKQALK---QLIDPNLATKMPKQVDDFQLSIFESEDKGDKYKNGY	262
DB	425	KOTIVVNPSENSITNAKIKVQAYHSSYNNIGQINKVDYD---IKIYQ-VPRGYTLNKG	480

```

QY 263 QNLSGLVPTKPPTPGDPMPNPQDTSVLIRKVAIGDYSK-----305
D 481 D-----VNTKELT-----DVTNQYLQKITGYGNNNSAVIDFGNADSAVYVMN 522
QY 306 -----LLEGATLQTLGD-----NV 319
D 523 TKFOYTNSESPPLVQMTLISSTGNKSVSTGNALGFTNNQSGAGQEQYKIGNYWMEDTNK 582
QY 320 NSFO-----ARVFSN---DIGERIELSDGTTLDEL-----NSPAGYS 355
D 583 NGVQELGEKGVNTVYVFDNNNTNKYGEAVTKEDGSLIPNLPNGDYRVEFSNLPKGYE 642
QY 356 IAE-----358
D 643 VTFSGQGNNEELDSGLSVITVNGKDNLSADLGITYPKYKYNLGDYWMEDTNKNGIOPDE 702
QY 359 -----PIITKVAAGKY-----TIDCK-----QIENPKEL-----VEPIYEAANDF 397
D 703 KGISGVYVTLKDENGVNLKVTITDADGKYKFTDIDNKNYKVEFTTPEGYPTPTVTSGSDI 762
QY 398 EEFSS-VLTTONY-----AKFYAKKNGSSOVVYCFNADLKSPDSED---GGKT 443
D 763 EKDSNGLITTVINGADNMTLDSGFTKPKYNLGNVYMEDTNKDCKO--DSTEGISGVT 820
QY 444 MPEFTTGEVYTHIAGRD-LFKYTVKPRDTPDPLKHLKVIENKGRKGAIEY---499
D 821 VTLKENGSEVILQFTKTDCKGYOFT-----GLENGTYKVEEETP 859
QY 500 SGLTEPQLRATQOLAT-YTFTDSALDKDK-----LKDY-----HGF 536
D 860 SGTPTPOVSGTDEGIDSNSTGTGAKDKDNPTIDSGFYKPTYNLGDYWMEDTNKNGVQ 919
QY 537 DANDSTLAVN-----KILVEYADSNP-QLTDLDFIPNNKYSGLTQWHPED 586
D 920 DKREKISGTVTLKENDKVLKVTITDEKGYOFTDL-----NNGYK-----963
QY 587 LVDIIRREDKEVIR-VTHNLTRK-----TVTGL---AGDRTKDFHEIEIKNNKQEL 636
D 964 ---VEETPPSGYPTSVTSNGNDEKDSNGLTGTGVIKADNMNLTDSGF---YKTPKYSL 1016
QY 637 -----LSQVTIDTNTLNEFKDKATINLKHGESLTLQGLPEGYSVLAKETSECKYK 689
D 1017 GDYVWYDSNKGQODSTEGIKDYKVTLLNEKEGEVIGTGTDEKGYCFDNLDSGKYKVI 1076
QY 690 VNSQEVANATVSKTGITSETTLAFENKKEPVPTGVQDKI-NGY 732
D 1077 F--EKPAGLTQGTGNTTEDDKADGGEVDYTIIDHDFTLIDNGY 1118

RESULT 2
US-10-615-383-10
; Sequence 10, Application US/10615383
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P0633503/BAS
; CURRENT FILING DATE: 2003-07-09
; PRIOR FILING DATE: 1999-01-25
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-615-383-10
Query Match 3.68; Score 142; DB 6; Length 930;

```

```

Best Local Similarity 18.6%; Pred. No. 0.039;
Matches 161; Conservative 131; Mismatches 295; Indels 278; Gaps 42;

QY 9 KINTLNTQR--VLSSKSK-----RFTVTLVGVPLMIFALVSMGAKTVEGL-----VES 56
D 2 KNNLTLTRKKRPINKSKYAIRKFTVTGTASI-----VIGALLFGLIGNEAKAE 51
QY 57 STFNALINPDSSSEVRWYGESEYVGRHPYKQFVRAHDLRVNLGSRSGYQYCFNLKAPF 116
D 52 NTYQDYK-DSNMD-----DELSDNSQSSNEKND---VNNQGSIN 89
QY 117 LGSDDSVYK-WYKHDGISKTFEDYAMS-----PRITGELNOKLRAYMNGH 163
D 90 TDDDNQIKRKREFTNSNDIAENRSKDITQSTNVNENEAFTLOKTPQDVTQKEVYVEPSS 149
QY 164 PQNANGIMEGLEPLNARIVTQEAIVWYSDAPISN-PDEFKRESBSNLVSTQSLMRQ 222
D 150 VESSNSSMDYAOQPSHTTINSIASIOTSDEMSRVSDFANSKTIESTSKREEMETIQ 209
QY 223 ALKQILDIPNLATKMPKQVPPDPLSTIPESBDKDGKYKQNLGSLGVPTKPPTPGDP 282
D 210 -----PKVREDSTISQPSISKINDEKISNODELLN---LPINEXKVRP 252
QY 283 MPN-OPQTSVLIRKVAIGDYSKLEGATLQTLGDVNSFGQARVSSNDIGERIELSD 340
D 253 LSTTSAPSSSKRYTVNQLAEQ-----GSNVHLIK--VTQDSITEGYDDSD 297
QY 341 GYTLTFLNLPAGYSAIEPTTFVEAGKYTI-IDGQI-----378
D 298 GIIRKHADEMLI-YDVFEEVDKVSQDITVINDIKNTVPSDLTDSFALPDKINSGEII 356
QY 379 -----ENPKIEIPEVS--VEAYNDEEFSVLTONYAKFYAKN--NGSSOVVYCFNA 429
D 357 ATGYTONTNKQIYTFITDYDYKYNKHAHLKLS-----YDKSKVPNNNTLVDYEKI 410
QY 430 DLKSPDSEDGKTWPDF-----TTGEVY--YTHIAGDLKRYKPRDTPDFELKH 481
D 411 ALSV-----NKTIVYEXKPNENRANLOSMTNI---DTKNHTV---EQTLYINP 456
QY 482 IKRVIKGRKGOALIEYSGLTE---TOLBAATQOLAIYFTTDSAEI-DKDKLDYHGF 537
D 457 LR-----YSAKTNVNISGNGEGSTIIDSTTIKRYKYGQDNOLPDSRITDYSEYED 510
QY 538 MNDSTLAVAKILVEYADSNPQTLDFEIPNNKYNQSLGTQWHPEDLVDIRMED--595
D 511 VTND-----DYAQ-----LGNNNVNINFGNIDSPY-IIVYISKYDPA 547
QY 596 KKEVIVPTHNLTLKRTVTLGAGD-----RTKDFHE 626
D 548 KDDYTTIQQVTVMQITINEYTGFEFRASYDNTIAFTSSGQGGDLPEPKYKIGDYVE 607
QY 627 -----TELKNNQOELL-----SOTVKTDKTNLEFKDKATINLKHG--ESI 665
D 608 DVDKDGIQNTNDEKPLSNVLTLYPDGTSKSVRFDSESKYQFDG-----LKNGLTYKI 662
QY 666 TLQGLEPGEYSVLAKET-----DSEGYKVVNSQEVANATV-----700
D 663 TFE-TPEGYTPPLKHSGTNPALDSEGSVWVTINGODMTIDSGFYQTPKYSIGNVYWD 721
QY 701 -SKTGITSD-----TLAFEN 715
D 722 TNKDGIQDDEKISGVKVTLKDEN 746

RESULT 3
US-10-617-320-3676
; Sequence 3676, Application US/10617320
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
NUMBER OF SEQUENCES: 5206

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/617/320
FILING DATE: 10-Jul-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107/433
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariñello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3676:
SEQUENCE CHARACTERISTICS:
LENGTH: 1769 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...1769
SEQUENCE DESCRIPTION: SEQ ID NO: 3676:
US-10-617-320-3676

Query Match
Best Local Similarity 3.4%; Score 134.5; DB 6; Length 1769;
Matches 161; Conservative 112; Mismatches 268; Indels 325; Gaps 43;

OY 145 RINGDELAK-LRAVM-----YNGHP-----QANGIMEGLEPLNAIVTQEA 187
D 330 KVTYDTIOSKVLKAVIDQAPPRVKEYTLNGHTLPQVOQFQVP-----INNHRITPE-V 383
OY 188 WYSDNAPISNDPSFRESSENLVS--TSOLIMROAL-----KOLIDPLAKMPQOV 240
D 384 TYKLINE--TTAETLMLKRDADALINAEVYKLOVNOHLFDVTKYVNHQVT--PEQK 439
OY 241 PDDFO-----LSIFESDEKDKYKNGYONLISGLVPTKPTPGDPPMPNPQ 288
D 440 IDDERKLLSISFLGNLVSVDQGTAKFD-----GATMGNNTHVSDDHIDVTNP 491
OY 289 QTTSVLRKTAIGDY--KLECATLQLT-----GDN-----VNSFOARVSSNDIG---- 333
D 492 MKD--LAKGYMYGVSTDKLAGVMSNSONSYGGSNDWTFRLTAYKETVGNANYGHISS 549
OY 334 -----ERIEISDGYTYTLE--LNSPAGYSIAE 358
D 550 EMQWEKAYKIVPEPTREKELPSAKVITTEDANADKKVMDGALAYISIMNPGQWEK 609
OY 359 PITEKVDAGKVTII--DKQIENP-----NKEIVEPYSEAVYND-- 396
D 610 DIT-----AYRIAMNFGSQANPFLMTLDGIIKINLHTDGLGQGVLLKGYGSEGHDSG 662
OY 397 -----FEESVLTQNYAKFY-----YAKNKGSSQVYVCFNAD 430

```

```

D 663 HLNADIGKRIQGVEDKTLIEK--AKYGAHGLIHVNASITYESKTFENKIL----- 714
OY 431 LKSPDSEDEGKTMPTPTTGEVYKTHIAGR-----DLFKYTVPRD-----T 473
D 715 RKNPDGSYSYGMNMLDQGINIDAAVYDLAHLGLAWMEDLKKLGGDLPIYDVVNGQSG 774
OY 474 DPTFLKHI--KATIEKGR--EKQA-----IESGLTEQLRAATQALAI 515
D 775 DNGAMATHVLAKEINKQWRAPEMGHGEYDSTFHMAADLTGTYINKGINSATREFI 834
OY 516 YFTDSA-----ELDKDKLKDYHGFQMDN-----STLAVALKIVEXAQ-- 554
D 835 RNHOKDAMVGDYRSTYGGAANPILLGYSMKDFEBOGRSDYNGVTNLFADVWTKFQH 894
OY 555 -----DSNPOLTDLDFIPNNKYQSILGTMHEDVLDIIMEDKREVI-----PV 602
D 895 FTYSKWEKGFVMTD-----NGSTY-----KMPDEMRELVADADNNKVVYVRKSNV 942
OY 603 THNLTUKTYV-----GLAGDRKDFHELELKNKQOL-- 636
D 943 NSPLYRRTVTLNGRVIODGSAYLTPMNMWANGKRLSTDKEKMYFMTQAGATTWTLP 1002
OY 637 -----LSQTVKTDKTNLEFKDKGATINLKHGSLTLQGLPEGYSLYKREND-- 682
D 1003 WAKSKVLYLKLTDGCKTEBEELTVKDGKITDILANDQPYL-----YSKQINPMS 1054
OY 683 -SEG---YKRVNSOEANATV---SKTGIT---SDPTLAFENKEPV----- 720
D 1055 WSEGMIHYLDGFNSGTLKHWIISGDASAEIVKSGANDMLRIGGNKEKVALTQKLG 1114
OY 721 -----VPTGYDQKINGLALIVTAG 740
D 1115 PMTKYAVYGVYDNRSNAKASITVNTG 1140

RESULT 4
US-10-333-002-19
Sequence 19, Application US/10333002
GENERAL INFORMATION:
APPLICANT: Adderson, Elisabeth
TITLE OF INVENTION: Group B Streptococcus Polypeptides Nucleic Acids and Therapeut
FILE REFERENCE: 2511-1-001 (SJ-0039)
CURRENT FILING DATE: 2003-07-08
PRIOR APPLICATION NUMBER: PCF/US01/24795
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 09/634,341
PRIOR FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 560
TYPE: PRT
ORGANISM: Streptococcus agalactiae
US-10-333-002-19

Query Match
Best Local Similarity 3.4%; Score 134; DB 6; Length 560;
Matches 103; Conservative 40; Mismatches 130; Indels 150; Gaps 26;

OY 386 VEPY-SVEAYNDFEESVLTQ--NYAKFYAKNKGSSQVYVCFA---DLKSPDSED 439
D 72 VNPYGSIRPN-----FSCAVYQDINLRAIM--RKAGYHIIISDAAGTGKRALDA-S 123
OY 440 GSKTMDPTTGEVYKTHIAGRDLFKYTVKVRDTPDTFLKHKKVLEKGYREGQALAY 499
D 124 GQQLQTSNEPDPDPSYDGDGSHALIR-----RPTMPD-----GYRFGW--WY 164
OY 500 SGLTETQIARATQALAIYFTDSALDKD-----KLKD-----YHGF 535

```

```

Db 165 NG-----KIYNPYDSIDDAHLADANKNITIKFVILIPVGDIKLEDSIKXNGN 212
QY 536 GDMNDSTLAVAKILVEYAQDSNPOLDLDFEIPNNKKY-----QSLIGTOW-HPEDLVDI 590
Db 213 GGRVNGNANV-----TQVETPRMELNSTTTIPENQYFTRTGYNLIG--WHHDKDLADT 264
QY 591 IRME-----DKKEVI-----PYTHNLILKRTVYGLAGRTKDFHFEIELKNN 632
Db 265 GVEFTAGOSIGIDNNLDATNLTLYAWQPKREYVGSKVYGLDEDEKTRDFLP-----N 318
QY 633 KQELLQYVKTQDNTNLEFKDKATINLKHGESLTLOGLPBGYSYLKKEIDSEGYKXVNS 692
Db 319 PSETLQO-----ENPPLRQO--TKEEK-----VPGTSTISIDEQAYDEKVSSEI 362
QY 693 QE-----VANATVSKTGTS-----DETLAFENKKEPVPTGVDOKI-----NG 731
Db 363 TENKLNATGEADKTYDATGLQSLVSGVDLSTFN-----TRIKQKVRLOKXVENDNN 415
QY 732 YLA 734
Db 416 FLA 418

RESULT 5
PCT-US02-37235-44
; Sequence 44, Application PC/TUS0237235
; GENERAL INFORMATION:
; APPLICANT: Handfield, Martin
; APPLICANT: Hillman, Jeffrey
; APPLICANT: Proculuske-Pox, Ann
; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigens
; FILE REFERENCE: MBH01-662A
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1531
; TYPE: PR
; ORGANISM: Actinobacillus actinomycetemcomitans
PCT-US02-37235-44

Query Match
3.3%; Score 131.5; DB 1; Length 1531;
Best Local Similarity 20.0%; Pred. No. 0.45;
Matches 179; Conservative 120; Mismatches 323; Indels 275; Gaps 47;

QY 39 ALVTSNMGAKTYVGLVESSTPNAINDSS-----SEYKRYGIESYVKGHPYKQRRVAH 92
Db 541 ANTTATNGSVILG--NSSITGSHPSVSVSATVNGYTGFTGVKESGHFVSI----- 594
QY 93 DLKVNLEGRSYOVYCFNLKKAFLPGSDSSVKKWKYKHDKIDKFEDVAMSPRITDELN 152
Db 595 -----GSKNFERQIKNV--AAGNVANANST-----DAVNGS--OLFVAVSRV---EOG 634
QY 153 OKIRAVMYNGHPONANGIMEGLEPIAIVRY--TOEAVWYSDNABISNDE--SEKRESE 208
Db 635 WQTSSEVNGNGONGAASATATIKPSNQVKLAGKNLAVQNGNINPFSTQENVTFINVTT 694
QY 209 SNLVSISQSLMRQALQO-----LIDPNLATKMPKQVPPDFOLSTFESEDKGDKYN--KGIO 263
Db 695 QDLTAIGNTVKNFVSQNGGTINMGNNRITGVAEGTQDD-----DAVNFKQK 742
QY 264 NLLSG-----LVPTKPTPGDDPMP-----NQPOTSVLIRKAYAGISYKLLBGA---TIO 313
Db 743 SLUGSASTEIVEKKAQAGDENLADISVANGKAGDMGAKTEVSKRAVOAARV 802
QY 314 LTGD-----NVNSFO-----ARVSSNDIGERIELS-----DGYTTLTE--LNSPAGYS 355
Db 803 VNGSAPINYNKTDVNGVDITYAVTFNGTEAKSIPLYTKANGSGDKVMDKGLNTINGNM 862
QY 356 IAEPI-----TFKYVAGKVVYTI-----IDGQIENPNKEIVEPEYVE 392

```

```

Db 863 TFSASVANDGVKYVDNLSTIKYEDGKAAYAGTPGTNGANGTGDGKDVATVKNVREALNNA 922
QY 393 AY-----NDFEES-----VLTONVAKFYAKKNGSSQYVCE 427
Db 923 AMTITASKSDGEVYVNAANSVKNRGDVTYTDAGKNITQIRDKKFSATADNVEFTSYTG 982
QY 428 MADLK--SPDSEDEGKTMPPDFTTGEVYKTHA-----GRDLFKYTKRPR-- 471
Db 983 NTKLTGNGVEITNGPKLTQSGVDAGKKITTNADVIAANSKDVANGGOLFETAKAKTT 1042
QY 472 -----DTPDPTLKH-----IKVIEKRYREKGAIE-----YSGITEQL-- 507
Db 1043 VEKGDNDNIQISTEADAGHINRYVALNPSSLVGPRTNGHITIDNGNGYITGLTNTSWTG 1102
QY 508 -----RAAT--QLAIV--FETDSAEI-----PKDKLADYH 533
Db 1103 APPTGAAATBEDQSLYDKKFKDNVSLGCGNGSTTEKSLSHNGIKRNINGGDSQKVVTS 1162
QY 534 GFQDMDSTLAVA--KILVEYAQDSNPOLDLDFEIPNNKKYQSLIGTQWHPEDLVDI 590
Db 1163 GSGD--DVATVDLAQTTRKNKIDNADKDLANITD-----NGKVVITALGA-----V 1205
QY 591 IRMEDEK--EVIPTVTHNLILKTY-----TGLADKTRKDFHFEIELK 630
Db 1206 VKAADSTITVYTDFTDNTGQKTYKIKANIPPEKTAAMAGNNTLEGDSANPEKVNK 1265
QY 631 NN-----KQELLQYVKTQDNTNLEFKDK--ATINLKHGESLTLOGL-----PE 672
Db 1266 DDLALQOKANGVYTGKDDSLKVGKRGSGVAINGKOG--SLANGKRGANPVITIKTQGA 1324
QY 673 GSYLVKEDSEGYKXVNSQEVANATVSKTGITSDETLAFENKKEPVPTGVDOKI 729
Db 1325 G-----VNETNPKD--RLMVNNDVATL-----KDGJL-----KFAQNGSTEVITKT--LNOKL 1368

RESULT 6
PCT-US03-22113-30
; Sequence 30, Application PC/TUS0322113
; GENERAL INFORMATION:
; APPLICANT: Seattle Genetics, Inc.
; TITLE OF INVENTION: SGA-72M, A CANCER ASSOCIATED ANTIGEN AND USES THEREOF
; FILE REFERENCE: 9632-061-228
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: 60/395,752
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1456
; TYPE: PR
; ORGANISM: Homo sapiens
PCT-US03-22113-30

Query Match
3.3%; Score 131; DB 1; Length 1456;
Best Local Similarity 20.1%; Pred. No. 0.45;
Matches 160; Conservative 126; Mismatches 262; Indels 248; Gaps 44;

QY 54 VESSTPN-----AINPSSSEYRWYGESYV--KGHPYKQ-----PRVAHDLRYNLE 99
Db 466 VADSLPSRBNITLPALOPPSN-----VGSERHAPLEKDPPLVRESNANKVM--SLKERMD 519
QY 100 GRSYOVY-----CFNLKKAFLPGSDSSVKKWKYKHDKIDKFEDVAMSPRITDELN 143
Db 520 EPNAEDYVYNPQFENIRKRWDELEANSNSNDNDKNITTTQOKSAPPNRQKHKEFSIDKIS 579
QY 144 PRITGD--ELNOKIRAVMYNGHPQ--NANGIMEGLE-----PINAIRVTQEAIVWYSDNAP 195
Db 580 GKNTHEAVLILBPKVMAAREMEKLNKSGIILQVLPDEITFPLPLR-----KTYQOLP 632
QY 196 ISNPDESEKRESNUNVTSQSLMRQALQKULIDPNLATKMPKQVDP-----DFOLSTFE 250
Db 633 ---GNSSKENVEKN-----TEGIYVVPKREKDYSEDEIOESIITK 671

```

[illegible]

```

Db 1293 RAtGVePIrIRWmKNGIdLkMDeKkFvSDNlGt--LRiEADKNDIGNNCvYtMDeAGS0 1391
QY 50 VEGlVEStPNAiNPDSSSErWMyGESyRGHPYyQOfRvAHDLRvNlEGSRsYoyYcF 109
Db 1352 MTHvADVDePILlIPStQTN-----MTAAVGD-----RV--ELKcYVEASpASvYWF 1397
QY 110 NLKKAfPLGSDSSsYKwYKKKHdGt-----StFEEDYASPRITtGELNOKLRvYvNNGHPQ 165
Db 1398 --RGAIGtDT--KGyVESDgTLvYtQSASvEd-----AtYtTCkAS 1436
QY 166 NANiMGLEPDLNAiR-----vToEAv-----wYSDNAP 195
Db 1437 NPACKAERANLoVYtIASPDiKDPDvYtOESiKESHPsLYCPvSNPLPQISWtLNDK-P 1495
QY 196 iSNDeSkFRSEsNLvStSOLsLMRQALkOLIDPNLAtKMPK--QvPDDfOLSiFESEd 253
Db 1496 LiDDKtSwKtSDDKRKLHtVFAKlTtDsgvYKcVARNAGeGSKSpQvEYvIPLMDeSKY 1555
QY 254 KGDKYnKGyONLsGGLvYtKPPtPGDPMPMPNOpTtSULtRKYtAGDYSKLLGAtLQ 313
Db 1556 KKKVFAK-----EGEEVtLlCCpVSGFP-----VQIIMwYV--DgTVVEPKKKYGAAtLS 1602
QY 314 LTGNvNSfQARvSSNDIGERiELSDStYtLTLENSAGStAEP-----ITfYvEAGV 369
Db 1603 NDGLTLH-----fDSvSVKQ-----EGNY-----HCvAOSKGNlIDIVElSVL 1641
QY 370 YtTiDgKQIENPKNEiVEPySvEAYvNDPEEFsYtLTONYAKfYyAKNKGSSQvYvCfNA 429
Db 1642 AvPLVG---EDDNLvEFLGKDiSLsCD-----LQTEsDDKtTYtWMSiNGS-----ES 1685
QY 430 D----LKSPDS-----EDGKtMPtPDTtGEVYKtYHlAGRDLEfkYtVAKPRDTP 475
Db 1686 DRPDvNOiPSGHRlYtIDAKRPENNGKYM-----CRvYNSAGK--AERTtLIDvLEP 1735
QY 476 DTFLKHlKkYIEKtYRGEGALIEYSGLEtPQLRAAtOLAItYtPDSAEtDKDLKdYHGf 535
Db 1736 PVF---VEEPfEANOklLG---NNPtlQCvYtGNPKtPYtIMKIDGNdvDkSWL----- 1783
QY 536 GDMNDStYAVNAKtILVEYAODS-----NPQLDfLDFEiPN-----NNKYOSLIGtOWH- 583
Db 1784 FDEStSLtRiKtLGKSQoISCTAENKAGtASRDfIONtAPltPKNGDOEtIFRESEt 1843
QY 584 -----PEDLvDIIMEDKEvIpytHN---tLKRt-----vTGLAG 617
Db 1844 ItLCPvStSGD-fQITWmKQGLPtLENDAlPtLNDtRLtILNANRHdEDiYtCVANNTAG 1902
QY 618 DRtDfHEfEiLk-----NNKQEL-----SOTvYtDKtNLEfKGG---KAtIN 658
Db 1903 QvSDfDfVvYvQvLEPRiKNAvVtLEINeEEtILtCDAGNtPtpRAKMDfONGDLPKEAVf 1962
QY 659 LKHESLtlQGLPrgYS--YLVKEMDSGcyVvKvNSQvANAtYSKtGtSDE----- 709
Db 1963 vNNHtYvVvNNvTYHrGVYKcYtATNKGAvKtINHVtRKtPRESEtSEtELtVNLtR 2022
QY 710 tLAEENKKEPVyPtGVOKiNGYtALtYtIAGiSLiGvG--ItHtIRiRKHD 757
Db 2023 SttLEcYvDDAIGvISWtVNGKRFLEtDvQvLlAGRfLHtYSAKDD 2072

RESULT 8
US-60-490-890-1651
: Sequence 1651, Application US/60490890
: GENERAL INFORMATION:
: APPLICANT: Li, Martha
: APPLICANT: Rupunow, Brent A.
: APPLICANT: Webster, Kevin R.
: APPLICANT: Jackson, Donald
: APPLICANT: Wong, Tai W.
: TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
: FILE REFERENCE: D0310 PSP
: CURRENT APPLICATION NUMBER: US/60/490,890
: CURRENT FILING DATE: 2003-07-29
: NUMBER OF SEQ ID NOS: 2779

```

```

RESULT 10
US-10-258-898A-2582
: Sequence 2582, Application US/10258898A
: GENERAL INFORMATION:
: APPLICANT: Hyseq Inc
: TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
: FILE REFERENCE: 784F1PCT
: CURRENT APPLICATION NUMBER: US/10/258, 898A
: CURRENT FILING DATE: 2002-10-29
: PRIOR APPLICATION NUMBER: US/09/488, 725
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US09/552,317

```

;; PRIOR FILING DATE: 2000-04-25  
;; PRIOR APPLICATION NUMBER: US09/598,042  
;; PRIOR FILING DATE: 2000-06-20  
;; PRIOR APPLICATION NUMBER: US09/620,312  
;; PRIOR FILING DATE: 2000-07-19  
;; PRIOR APPLICATION NUMBER: US09/653,450  
;; PRIOR FILING DATE: 2000-08-31  
;; PRIOR APPLICATION NUMBER: US09/662,191  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: US09/693,036  
;; PRIOR FILING DATE: 2000-10-19  
;; PRIOR APPLICATION NUMBER: US09/727,344  
;; PRIOR FILING DATE: 2000-11-29  
;; NUMBER OF SEQ ID NOS: 7143  
;; SOFTWARE: pc\_files\_versions 1.0  
;; SEQ ID NO 2582  
;; LENGTH: 993  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-258-898A-2582

Query Match 3.2%; Score 128; DB 6; Length 993;

Best Local Similarity 20.1%; Pred. No. 0.43;  
Matches 145; Conservative 89; Mismatches 232; Indels 256; Gaps 35;

OY 85 YKORVAVHDLR--VNLG-GSRSYOYCCNLKAPL---GSDSSYKKKKKHGISTKFE 138  
DB 22 FKLRLVHLKLGFPNIESGAVWKQY---ODLGIPVLSAAGYVNCVKYK---LYGFE 75  
OY 139 DYASPRITGDELNOKLRAVYNGHPONANGIMEGLEPLNIRVTOEAVWYSDNAPISN 198  
DB 76 EYCSNPNI---EFGMALPEKYVNNKCKCEVYK-----IKYKE-----N 113  
OY 199 PDE---SEKRESENLVSTQSLSLMKQ-ALKQILDNPILATKMPKQVDDFOLISEEDKG 255  
DB 114 ETEIOEIKMEERNIIPREKPIEDEIERKENIKRISGSK----- 153  
OY 256 DKYKGYONLSSGGLVPRKPPPTGDPMPRPNQPTSVLIRKAYAGDYSKILEGATLOLT 315  
DB 154 -----KNLES-----IPHSDEKEVNIK-----P 175  
OY 316 GDVNSFQARVSSNDIGERIELSDGYTLTLELNSPAGYSIAEPTTFVYKAGKYTTIIDG 375  
DB 176 EDNEN-----LDDKDDDTTRVD-----ESINIKVEAEE---KA 206  
OY 376 KOIENPNK-----EIVEPYSEAVNDEEFSVLTTONYAKFYAKKNKG 419  
DB 207 KSCDETKEDEDEDEAEAESEEEDEDEDODNNEEEFECEYPPGAKVQYRGGRKN- 265  
OY 420 SSOVYVCNADLKSPDSEDDGKTMPTPTTGEVYK-THIAGROLFKY-----TYKP 470  
DB 266 --QMY--EASIK--DSDEG-----GEVLYLVHYCGWNV-RVDEMIKADKIYRP 308  
OY 471 RDTPTDFLKHIKKIVIEKGYREKGOAIEYSGLTETQOLRAATQALAIYFTDSAEIDKDKLK 530  
DB 309 ADKAVPK-IKHKRIKIKKLDKEKDEKISP-KNKKLR-----RLSKPPFQ 352  
OY 531 DYHGFDMNSTLAVAKLIVEYADSNPQLTDLDEFIPNNKKYOSLIGTGMHPEDLVDI 590  
DB 353 -----TNPSPEMWSKIDLTDAKNSDTAHKISEI-----TSILNGLOASSSAD 397  
OY 591 IRMEDKKKEIVYTNHNLTKRTVTGLAGDRTKDFHFEIELKNNKOLLQYTKDTKTNLEF 650  
DB 398 SEQDEBERGADMDNN-----GKEESKIDH---LTNNRNDLISKESONSSSLLLE 443  
OY 651 KDGRATILKKGESITLLOGLPEGYSYLAKETDSEGYKVKVNSQEVAAVTS----- 701  
DB 444 NKVHADLVYSKRPVSKPERLKRKIDLEVLSEDDYEDDEYTKRKRDVKKPTTSSKSPQIKR 503  
OY 702 -----KTG-----ITSDETLAPENN-----KEPVVPTGVDOKI 729  
DB 504 GKRRYCNTIECLTKGSPGKKEEKAANKESLCMENSSSSSDEDEETRAKMTPT---KKY 560

OY 730 NG 731  
DB 561 NG 562

RESULT 11

US-10-408-765A-2576  
Sequence 2576, Application US/10408765A

GENERAL INFORMATION:

APPLICANT: Ghosh, Soumitra S.

APPLICANT: Fahy, Erin D.

APPLICANT: Zhang, Bing

APPLICANT: Gibson, Bradford W.

APPLICANT: Taylor, Steven W.

APPLICANT: Warnock, Dale E.

TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

FILE REFERENCE: 660088.465

CURRENT FILING DATE: 2003-04-04

NUMBER OF SEQ ID NOS: 3077

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2576

LENGTH: 1090

TYPE: PRT

ORGANISM: Homo sapiens

US-10-408-765A-2576

Query Match 3.2%; Score 128; DB 6; Length 1090;

Best Local Similarity 18.3%; Pred. No. 0.49;  
Matches 153; Conservative 125; Mismatches 280; Indels 278; Gaps 42;

OY 41 VTSWAGATVEGLVESSTPMAINDSSSEYRMV---GYEYVRGHPYKQFRVAHDLRV 96  
DB 311 VFAPYKRSVLLIETIYTN--NSVALSHFEYAGAASEFFSQGP-----I 356  
OY 97 NLEGRSRYOYCCNLKAPLGLSDSSYK---WYKHHDISITKEEDYAMSPTITGDEL 151  
DB 357 DISTRKQHLIF--KEENP--NNDITVTKGSTANENTTPSPAPVAVAKATISARDPKL 413  
OY 152 NOKLRAVYNGHPONANGIMEGLEPLNIRVTOEAVWYSDNAPISNDESEFKRESENL 211  
DB 414 KTNLADAL--HSQS-----IIPLEIVNYIPD-----DNLSTISMDTK-EANNRSI 457  
OY 212 VSTQSLSLMKQALQILDNPILATKMPKQVDDFOLISEEDKGDKYKGYONLSSGL- 270  
DB 458 IFTS---MNEPTSTVVDSTVPSGYSVSDFMV---PGBD--NVKSRGFSQILRTMML 508  
OY 271 ---VPTKPPPTGDPMPRPNQPTTSVILIRKAYAGDYSKILEGATL---QLTGDVNSFQ 323  
DB 509 LLDLSLNENPETPG--LMDNSFSESSI-----TRSEELERATEYVGAINGVPERKQ 558  
OY 324 ARVFNSSN---DIGERIELSDGYTLTLELNSPAGYSIA---EPTTFVYKAGKYTT- 371  
DB 559 VITQTNHSHYPLANQVPIESTPSNNPTDSEINSPFVNQNTPOINTETETDEITSQR 618  
OY 372 -----IIDGKQ----- 377  
DB 619 AIEKAFSOLIKESQNSVYKIRKIANLSDQEDPATMELKSHNTTVLIPAIWKIPSATKL 678  
OY 378 --IENPKNEIYEPYV-EAYNDFE-----FSVLTTONYAKFYAKKNSSQVY 425  
DB 679 NLLSDDSDVDDEEFIMSEAGNSFSDTSKLDKDSFTVAFDDTSFQSQAAPNLIKISSYS 738  
OY 426 CFNADLKSPDSEDDGKTMPTPTTGEVYK-THIAGROLFKY-----TYKP 470  
DB 739 TSTLSEARPNDDSSAFHIRVSNIGRSSTNTPARLHGLQWKPRLPSDDTISMDAPYIEL 798  
OY 454 KYTHIAGDLFR--YVKKPRDT-----DDPTFLKHIKKIVIEKGYR 491  
DB 799 TPVITMGKIDITKIDYSFLPRETITPSDQVILLPTITSENSMASTFENIKF--HKDFI 856





```

Db 783 SPLAKIAPVITGNSKLEGRFF-----GKGIESHNTSADKREILAPFVRD-----ETF 834
Qy 401 S-----VLTONAKFYAK-----NKGSSQYVVCFNADLKSPD- 436
Db 835 GNTALLKKAESGECOLSTONLIQMAEDSHLDPTSQLSRKGSF-----GDVASPPD 887
Qy 437 -----SEDCGKMP-----DFTT-----GEVYTHIAGD 462
Db 888 MLFPQDAHLVQARVHSQTEISEVEKVIILPPRPVLNDVSAALQCLGCEWLSLPAGRE 947
Qy 463 LEKTYVRPDPPTFLKHKIKVIEKGRKGOALEYSGLTETOLRAATOLAIFYTTDSA 522
Db 948 -----VGRGEVNP-EFPAVQPV-----SPLNPGVISPMAFMOTI 983
Qy 523 ELDKDKLDYHFGDMNDSTLAVAKILVEYADSNPOLTDLDFIPNNKTKOSLIG-- 580
Db 984 VPDR---KDFYSSNVVPDKTHEVGSYL---AAMSPSDOT-----LSSPASIVAYG 1029
Qy 581 QMHPEDLVIDIRMEDKKEVIVPTHNLTRKTVTGLAGDRTKDF--HFETELKN-KKOELL 637
Db 1030 KGLQVEVEIYR---ETIVQKSEFLERSAGLEKLEKETETFPSEKYESDIGNLSPSKLI 1086
Qy 638 SQYVKTDTNLEF--KDGKATINLKHGSLTLQGLPEGYSYLVKE-TDSEGYKRVNSOE 694
Db 1087 GSTEPRRATSECHPEELKEKTEYKAEAPLITESAFDAGFEKLEKITEITAPPYQPVSYRE 1146
Qy 695 VANATVSKTGITSDFT 710
Db 1147 ---ETHEKESQSEOT 1159

```

## RESULT 14

```

US-09-200-650E-3
; Sequence 3, Application US/09200650E
; GENERAL INFORMATION:
; APPLICANT: Patil, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eichlin, Deirdre M.
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-3

```

Query Match 3.2%; Score 125.5; DB 5; Length 930;

Best Local Similarity 19.4%; Pred. No. 0.59;

Matches 115; Conservative 71; Mismatches 215; Indels 193; Gaps 26;

```

Qy 183 TOEAVWYSDNAPISNPDESFRSESNLVSTLSLKRQALQKQIDNLAIRKMPQVDP 242
Db 86 TOTAT---ADQPKYMSASATYKETSMMQSPONKATNOS-----TTKTSTNVTN 132
Qy 243 DFOLSIFESEDKGDKYNGYONLSGLVPTKPPPGDPMPNPOTSTVYLIRKTAIGD 302
Db 133 DKSSITYSNET--DK-----SVLQAKDVSTTPKTTTIKPRILNMAVTVAAQ----- 180
Qy 303 YSKLLEGATLQLTGNVNSFQARV-FSSNDIGERIELSDGYTTLT-----LNSPAG 353
Db 181 -----QGTNVND---KVHFSNIDIALDKGHVNOQTGKTEFNATSSDVKLKAN 225
Qy 354 YSIAPFI-----TFK-----VEAGKVYITIIQKQIENPNKELVEVYSVEAVYNDPEFSLV 403

```

```

Db 226 YTIIDSVAGEDEFTFFKYGQYFRPGSVRLPSQONLYNNOGNI---AKGITDSTNTTYY 282
Qy 404 TTQYAKFY-----YAKNKGSS-QVYCFNADLKSPDSED-----GKTMTP 446
Db 283 TETNYVDQYTVNRGSEFQVAFARKNATTDKTAYMKEVTLGNDTYSSEIIVDYGNKKAOP 342
Qy 447 DETTEGVKTYTHAGDLKRYTKPRDT-DPOTFLKHKKVIEKGRKGOALEYSGLTET 505
Db 343 LIISTNYINNEDESNNMTRAYVNOPTKTYTKOTFEVNL-----TGKFNPMANFK----- 392
Qy 506 QLRATOLAIFYFTBSAELDKDKLDYHFGDMNDSTLAVAKILVEYADSNPOLTDL 565
Db 393 -----IYEVDQNOF-----VD 404
Qy 566 FEIPNNKYSQSLIGTQWHPEDLVIDIRMEDKKEVIVPTHNLTRKTVTGLAGDRTKDFH 625
Db 405 SETPDTSKLDV-----TDQFDVYISNDK-----TATVDLMKGOT----- 440
Qy 626 ETELKNNKOELLISO-----TYKTDKTNLEFPDGKATINLKHGSLTLQGL 669
Db 441 -----SSNKQYITIQVAYVDPNSSTDNCKIDYITDPTKTYSMSNSYSNVN---GSSITANG 492
Qy 670 LPEGYS---YLVKETDSEGYKRVNSQEVANATV---SKTGITSDETLAFENNK 717
Db 493 DQKKYNLDYWDIEDINKKG-KODANEKGIKGYVILKDSNGKELDRTTDDENGK 545

```

## RESULT 15

```

US-10-617-320-3251
; Sequence 3251, Application US/10617320
; GENERAL INFORMATION:

```

```

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS

```

```

; NUMBER OF SEQUENCES: 5206

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESS: 100 Beaver Street

```

```

; CITY: Waltham

```

```

; STATE: Massachusetts

```

```

; COUNTRY: USA

```

```

; ZIP: 02354

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: CD-ROM ISO9660

```

```

; OPERATING SYSTEM: <Unknown>

```

```

; SOFTWARE: <Unknown>

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/10/617,320

```

```

; FILING DATE: 10-Jul-2003

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/107,433

```

```

; FILING DATE: 30-Jun-1998

```

```

; APPLICATION NUMBER: 60/085131

```

```

; FILING DATE: May 12, 1998

```

```

; APPLICATION NUMBER: 60/051553

```

```

; FILING DATE: July 2, 1997

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Ariniello, Pamela Deneke

```

```

; REGISTRATION NUMBER: 40,489

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (781)893-5007

```

```

; TELEFAX: (781)893-8277

```

```

; INFORMATION FOR SEQ ID NO: 3251:

```

```

; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1972 amino acids

```

```

; TYPE: amino acid

```

```

; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein

```

```

; HYPOTHEICAL: YES

```

ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1972  
SEQUENCE DESCRIPTION: SEQ ID NO: 3251:  
US-10-617-320-3251

Query Match 3.2%; Score 125.5; DB 6; Length 1972;  
Best Local Similarity 19.5%; Pred. No. 1.7;  
Matches 153; Conservative 120; Mismatches 297; Indels 213; Gaps 37;

QY 9 KLNTLNTQRYL---SKNSKRFYV---TLVGFLMIFALYTSWGAFTVGLVESSTPNA 61  
Db KINSQNLAEPLPNTGSKNEROALVAGASLALGILIFAVSKKKYKNTVHLV----- 145  
QY 62 INPDSSSEYRWYGESYVRGHPYKQPRVAHDLRVNLEGSRSYQVYCPNLKAPPLGSDS 121  
Db 146 -----LVAG-----IGVGLVSVHALENHLLNNTDYELTSGEKL 181  
QY 122 SVKKWYKHH-----DGISTKFEDYAMSPRTGDELNOKLRAYMYN-----GHPQAN 168  
Db 182 PLKREISGYTYIGIKGKTT---SFEVSNQEKSAAPTKQKVDYNTVNFVDHPSTVQ 239  
QY 169 GIMEGLEPLNAIRVTOEAWYYSNDAPISNPDESFKRESNLYSTSQLSLMRQALKOLI 228  
Db 240 AIOE-QTPVASTKPTVEYVEKPESTELINPRKEKQSSDQ-----EQLA 284  
QY 229 D-PLATKMPKQYDDQOLSI FESEDDGDKYKNGYQNLGGLVPTKPTPGDPPMPNQ 287  
Db 285 EHKMLETKKEK-----ISPKEKTGVTLNPQDEVLSGQL-----NK 321  
QY 288 PQTTSVLIRKAYIG---DYSKLEGATLOLTGDNVNSFOARVFSSNDIGERIELSDGYT 344  
Db 322 PE---LLYRETEIFKIDPOE-----EIQENPDIVEGI-VRKQEGKLGKVEI---VR 368  
QY 345 LTELNSP-----AGYSIAPETFEKVEAGVYTIIDGKOIEN--PNKEI-----VE 387  
Db 369 IFSVKNKEVAREIYSTTAPRIPIVEKGTAKTQVKEQPEVGEHKGQGAIVERPAIQ 428  
QY 388 PYSEVANDEFEESVLTQNYAKFYAKNKGSSQVYCFENADLKSPPDSEDDGKTMFPD 447  
Db 429 PELPEA-----VSDKGVEYQPALSK-----AVITDKGETEVQPESPD---TVSD 472  
QY 448 FTTGEVAYTHIAGRDLEKFTYVKKPRDTPDPTFLKHKKIVIEKGYREKGOATEY----- 499  
Db 473 --KGEPEQV---APLEPKYGNIEQVAKPET---PVEKTKEOG-PEKTEEVYVAKTEETPV 522  
QY 500 ---SGLTE-TQLRAA---TOLAITYFTDSALDKLIDYHGFQDMN--DSTLAVAKILV 550  
Db 523 NPNEGTEGTISQEAENPVQPAEESTJNSEKVSPTTSSEWNGEBSNPDSSTSVG---- 578  
QY 551 EYADDSNPQTLTDLDFIPNNKYSQSLIGTOWHPED--LYDIIEMEDKKEVYIPV----TH 604  
Db 579 -----ESNKKPEHND---SKNENSEKIVEEYVNPNEGTEGTSNQETEKQVQPAELQTN 630  
QY 605 NLTLRKVTYGLAGRTKDFHFEIELKNNKOELLQOTYKTDKTNLEFKDGK----- 654  
Db 631 SGRKANENTGEVSNKRPDSKRPVDESNOPEKNGATKPENSGNTTSNGQTEPEPKLELR 690  
QY 655 -----ATINKHGESLTLQGLPEGYSYLVKRETDESGYKVKVNSQEVANATVSKTGIT 706  
Db 691 NWSDIELYSQTNCTYRQHVSLDGIPE-----NTDYTFVVKVKSAPKDVYIPVASIT 741  
QY 707 SDE 709  
Db 742 EEK 744

Search completed: August 19, 2003, 15:29:15  
Job time : 40 secs